UDC 681.518.22:606

A method for the assessment of the parameters of an automated system for analysing the microbiological objects motility

Yu. E. Megel[1], A. I. Rybalka[2], G. P. Venkov[3], S. V. Kovalenko[4], I. V. Chaly[1]

[1] Kharkiv Petro Vasylenko National Technical University of Agriculture, megelye@gmail.com
[2] Kharkiv National University of Radio Electronics, (Kharkiv, Ukraine)
[3] Technical University of Sofia,(Sofia, Republic of Bulgaria)
[4] National Technical University "Kharkiv Polytechnic Institute" (Kharkiv, Ukraine),

A large number of biological objects are characterized by their movement. Dependence of the movement direction of individual cells or its line by using of external light, electromagnetic radiation or chemical effects on a solution can be examined by using optoelectronic system for tracks analysing and counting the number of objects. In the case of analysing such objects characteristics, their number in the system field of vision is a random number. The similar problem arises in the biotechnology of reproduction and selection of livestock when creating new breeds with programmed properties and when applying artificial insemination.

In the process of creation such systems the problem of determination the probability with which the system can distinguish two or more objects appearing in its field of vision is arisen. In the paper the method of determination the probability depending on the system resolution and objects number in its field of vision is suggested.

Since the objects number in the field of vision can be varied or set by changing the solution concentration, the goal of the work is to determine the number of objects when the object merge does not occur with a probability that exceed some predetermined level of significance that is close enough to one.

Keywords*: microbiological object, optoelectronic system, resolution, field of vision, probability.*

1. Introduction. One of the important problems in the screening of new substances is the problem of assessing the impact of these substances to biological objects. Nowadays it is used the method of biological tests according to reaction of microorganisms to chemicals by measuring the motility characteristics of a large number of biological objects [1]. It requires the creation of modern automated optoelectronic systems. Similar problems are in the development of such systems for biotechnology of reproduction and breeding of livestock in creation samples of animals with programmed properties, e.g., for assessing the sperm quality of its motility [2] or in electro manipulation of cells.

Such an approach implies a set of methods for the action of pulsed and harmonic electric fields on living cells by means of specialized equipment with the need for reversible or irreversible electrical breakdown of membranes (electroporation, electrofusion, electrolysis), convergence before breakdown, sorting (HF dielectrophoresis), increasing the permeability of membranes (electropermeabilization), ensuring the transfer and integration of genes, DNA, RNA, viral macromolecules into the cell through the membrane (electrotransfection) [4-7] etc. In the process of developing such systems it is necessary to evaluate the upper limit of observed objects number because the system technical characteristics, such as system field of vision and its resolution, depend on their quantity, and, accordingly, the accuracy of the system as a whole [3].

2. Problem statement. The problem of object number assessing by means of optoelectronic system can be written in the following formulation. In the system field of vision n objects appear independently of each other. If the distance between objects is less than a value δ – the system resolution, then two objects are taken as one, if the distance is greater than δ, then the objects are taken as separate ones. It should be noted that three, four etc. objects can be merged according this rule. Since the object coordinates are unknown and their appearance is random, an event in which a merger of two or more objects occurs or does not occur is random. At the same time, their probabilities depend on the system resolution δ and on the quantity (n-number) of appearing objects in the system field of vision [3, 7].

It should be taken into account that the system parameters have technical constraints and the number of microbiological objects in some cases can be varied or set. The optimal situation is a case, when the merger of any two objects does not occur at all.

Therefore, we can consider a question at what quantity of objects n the above-mentioned event has a probability exceeding some predetermined level of significance that is close enough to one.

The problem identified above can be formally formulated as follows: there are n pairwise independent random vectors $\overline{\xi}_1, \overline{\xi}_2, ..., \overline{\xi}_n$, equally distributed with uniform law over some region $V \subset E_k$ (where E_k – k – dimensional Euclidean space, that has a dimension equal to two for optoelectronic system) [3, 8].

3. Calculating the probability of non-merging objects. Consider an event in which there are following inequalities:

$$\rho(\overline{\xi}_i, \overline{\xi}_j) > \delta,\ i \neq j,\quad i, j = 1..n \tag{1}$$

where $\rho(\overline{\xi}_i, \overline{\xi}_j)$ is the designation of the Euclidean metric in E_k. The probability of this event we will denote as $P_v(n, \delta)$. It is necessary to find the maximum value of n when

$$P_v(n, \delta) \geq \alpha. \tag{2}$$

Apply an inductive approach to solving this problem. The probability $P_v(n, \delta)$ can be recurrently expressed through $P_v(n-1, \delta)$, what, by taking into account an equality $P_v(1, \delta) = 1$, allows to consistently find $P_v(n, \delta)$. This approach will be realized below when $k = 1$ and $k = 2$ (cases, which are important for practical using).

Preliminarily get the general recurrence relation. For this purpose we partition the region V into parallelepipeds Π_i in phase space E_k with axes $z_1, ..., z_k$. In each parallelepiped we fix a point z_i, the Lebesgue measure of each parallelepiped is denoted as $\mu(\Pi_i)$, further the Lebesgue measure (volume) of the corresponding set we will denote as symbol $\mu(\cdot)$. Introduce the events: A – when a distance between any pair is greater then σ and the set of hypotheses $\{H_i\}_{i=1}^{S}$, where s – quantity of partition sections (number of parallelepipeds), and i^{th} hypothesis H_i correspond to the event when the first random vector ξ_1 gets into i^{th} parallelepiped.

In this case $P(A) = P_v(n, \delta)$. Find probabilities $P(H_i)$ of hypotheses and conditional probabilities $P(A / H_i)$ of our event under condition that an i^{th} hypothesis is true. Since a random vector ξ_1 is uniformly distributed over V we get

$$P(H_i) = \mu(\Pi_i) / \mu(V) \tag{3}$$

Now assume that hypothesis H_i takes place. In this case an event A occurs when following conditions simultaneously fulfil in δ – neighbourhood of a point z_i: $Uz_i(\delta)$ except the point ξ_1 any other points do not locate, i.e. $\overline{\xi}_2, ..., \overline{\xi}_n$ lay in a region $V \setminus Uz_i(\delta)$ (symbol \ – is a set difference symbol) and in a region $V \setminus Uz_i(\delta)$ random vectors $\overline{\xi}_2, ..., \overline{\xi}_n$ occupy positions in which $\rho(\overline{\xi}_i, \overline{\xi}_j) > \delta$, if $i \neq j$ and $i, j = 2...$. Probability of this event in our designations is equal to

$$P(A / H_i) = \left[\mu\left(V \setminus Uz_i(\delta)\right) / \mu(V)\right]^{n-1} \times \times P\left(V \setminus Uz_i(\delta)(n-1, \delta)\right) \tag{4}$$

By the formula of total probability [3,10] we get:

$$\begin{aligned} &P_v(n, \delta) = \sum_{i=1}^{S} \mu(\Pi_i) / \mu(V) \cdot \\ &\left[\mu\left(V \setminus Uz_i(\delta)\right) / \mu(V)\right]^{n-1} PV \setminus Uz_i(\delta)(n-1, \delta) = \\ &= 1 / \mu^n(V) \sum_{i=1}^{S} \mu^{n-1}\left(V \setminus Uz_i(\delta)\right) \times \\ &\times PV \setminus Uz_i(\delta)(n-1, \delta)\mu(\Pi_i) \end{aligned} \tag{5}$$

In the formula (5) implement a limit transition, letting the partition sections number s tend to infinity, and the Lebesgue measure of i^{th} section tend to zero. Since the expression under sign $\sum$ in the equality (5) present a Darboux sum for function $f(z) = \mu^{n-1}\left(V \setminus U_z(\delta)\right) PV \setminus U_z(\delta)(n-1, \delta)$, then in the limit we get the final recurrence relation

$$P_v(n, \delta) = 1 \setminus \mu^2(V) \int_{...} \int_{W} V \mu^{n-1}\left(V \setminus U_z(\delta)\right) \times \times PV \setminus U_z(\delta)(n-1, \delta) dz \tag{6}$$

where $z = (z_1, ..., z_k),\ dz = dz_1 ... dz_k$.

Thus, the relation (6) let reduce the problem for n points into the problem for $n-1$ point. However, further analytic research is hampered on the one hand by the arbitrariness of the dimension of the phase space k and the set E_k. On the other hand, there are practically important situations when $k = 1$ or $k = 2$ and $V = [0,1]$ or $V = [0,1] \times [0,1]$, i.e. the

field of vision is a thin enough band which, without restriction of generality, can be considered as of unit length or as a unit square in the plane. Values obtained for n in this case, for $k=1$ have linear density and for $k=2$ have planar density at which the equipment resolution with a high enough probability allows to distinguish all objects. Furthermore, it can practically be considered that there is a power link between these densities, i.e. the planar density is a square of linear one, so the planar variant can be reduced to linear one (just like any other, however). Therefore, in the following calculations our assumption when $k=1$ and $V=[0,1]$ is true.

Then, the objective means, that n points are uniformly thrown on the interval $[0,1]$ and probability $P_{[0,1]}(n,\delta)$ that the distance between any two of them was not less than δ should be found. It should be noted that in the case when $n=2$ this is the formulation of the problem, which is known in probability theory as "Probability of meeting" [3, 8, 9].

Based on the assumptions made by us, the Lebesgue measure is generally a length, so we will denote it as symbol $|*|$. Next if $V=[0,1]$, then $|V|=1$, so the equality (6) will take the following form

$$P_{[0,1]}(n,\delta)= \\ =\int_0^1 \left|[0,1]\setminus Uz(\delta)\right|^{n-1} P_{[0,1]\setminus Uz(\delta)}(n-1,\delta)\,dz \qquad (7)$$

From (7) it is clear that by decreasing the point number by 1, we change the original set [0,1] to some other one. It would be advisable that this change lead again to an interval (in this case the probabilities participating in the recurrence relation would be of the same type), however it does not occur. Further we will show that all of this can be reduced into an interval nevertheless, however its length will not be one. Therefore, we will solve the problem for a more general case, when $V=[a,b]$. Then an equal (6) takes the form:

$$P_{[a,b]}(n,\delta)=1/(b-a)^n \times \\ \times\int_a^b \left|[a,b]\setminus Uz(\delta)\right|^{n-1} P_{[a,b]\setminus Uz(\delta)}(n-1,\delta)\,dz \qquad (8)$$

Look in more detail the structure of a set $[a,b]\setminus U_z(\delta)$. On a line the closed δ – neighbourhood of a point z is an interval $[z-\delta,z+\delta]$. It is clear, that if its radius $\delta\ge(b-a)$, then $P_{[a,b]}(n,\delta)=0$ when $n\ge 2$ (on such an interval it is impossible to put even two points, separated from each other by more than δ).

Assume that $\delta<b-a\le 2\delta$, i.e. the radius of a neighbourhood is less than the length of an interval, however the dimension of a neighbourhood is greater than the length of an interval. In this case it is clear, that it is impossible to put more than three points in fulfilling our requirement, i.e. $P_{[a,b]}(n,\delta)=0$ when $n\ge 3$. And when $n=2$, the situation is similar to the formulation of "Probability of meeting problem"[10]. The desired probability can be get by using the relation (8) or as a geometric probability [3, 9] and it will be equal to

$$P_{[a,b]}(2,\delta)=(b-a-\delta)^2/(b-a)^2 \qquad (9)$$

The most general situation is when $b-a>2\delta$ (it should be noted that in practical cases $b-a$ is much bigger than 2δ). Then on interval $[a,b]$ it is possible to distinguish three zones of location of the point z, for two of which $[a,b]\setminus U_z(\delta)$ is an interval, and for the third it is a union of two intervals.

Thus, we get:

$$[a,b]\setminus U_z(\delta)=\begin{cases}[z+\delta,b],\ if\ z\in[a,a+\delta];\\ [a,z-\delta],\ if\ z\in[b-\delta,b];\\ [a,z-\delta]\cup[z+\delta,b],\\ \qquad if\ z\in[a+\delta,b-\delta].\end{cases} \qquad (10)$$

Therefore, the form of the integrand depends on the region of hitting z into the above-mentioned zones of an interval $[a,b]$. Hence based on (10) it is possible to write:

$$\int_a^b \left|[a,b]\setminus U_z(\delta)\right|^{n-1} P_{[a,b]\setminus U_z(\delta)}(n-1,\delta)\,dz= \\ =\int_a^{a+b}(b-z-\delta)^{n-1}P_{[z+\delta,b]}(n-1,\delta)\,dz+ \\ +\int_{b-\delta}^{b}(z-\delta-a)^{n-1}P_{[a,z-\delta]}(n-1,\delta)\,dz+ \\ +\int_{a+\delta}^{b-\delta}(b-a-2\delta)^{n-1}P_{[a,z-\delta]\cup[z+\delta,b]}(n-1,\delta)\,dz. \qquad (11)$$

Prove that the first two summands are equal. For this it should be noted at first that $P_{[a,b]}(n,\delta)$ depends only on $n,\ \delta,\ (b-a)$ and does not depend on the interval endpoints. Therefore for any a and b the following equation occurs

$$P_{[a,b]}(n,\delta)=P_{[0,b-a]}(n,\delta). \qquad (12)$$

Then in the second integral of the equality (12) we make a change of variable $z = a + b - u$, then

$$\int_{b-\delta}^{b} (z-\delta-a)^{n-1} P_{[a,z-\delta]}(n-1,\delta)\, dz = \\ = -\int_{a-\delta}^{a} (b-u-\delta)^{n-1} P_{[a,b+a-u-\delta]}(n-1,\delta)du \quad (13)$$

By substituting u back in z and by considering (13), we get:

$$\int_{b-\delta}^{b} (z-\delta-a)^{n-1} P_{[a,z-\delta]}(n-1,\delta)dz = \\ = \int_{a}^{a+\delta} (b-z-\delta)^{n-1} P_{[z+\delta,b]}(n-1,\delta)dz \quad (14)$$

Then we will consider the third summand of the equality (11). Here is the above-mentioned situation. Calculating the probability $P_{[a,b]}(n,\delta)$ is reduced to finding the analogous probability for $n-1$ points, only the set V changed its structure and became the union of two intervals. The further using of the recurrence relation in this form will lead to an even more complex transformation of V, etc. Naturally, it is inconvenient for practical using. Therefore we research in more details the probability $P_{[a,b]\cup[c,d]}(n,\delta)$, where $a<b<c<d$. Show that it can be expressed through probabilities $P_{[a,b]}(\cdot,\delta)$ and $P_{[c,d]}(\cdot,\delta)$.

Really, if $V=[a,b]\cup[c,d]$ and on this set n points are randomly selected, then it is possible in the case when k points hit on $[a,b]$, and $n-k$ points hit on $[c,d]$. And k changes from 0 to n. Since the choice of k points from n can be carried out by C_n^k ways and probabilities of hitting k points on interval $[a,b]$ and $n-k$ on interval $[c,d]$ will be equal to $\left[\frac{(b-a)}{(d-c)+(b-a)}\right]^k, \left[\frac{(d-c)}{(d-c)+(b-a)}\right]^{n-k}$, then the probability of such distribution of points on intervals is equal to $\frac{C_n^k (b-a)^k (d-c)^{n-k}}{[(d-c)+(b-a)]^n}$. If the above-mentioned event has occurred, then under this condition the probability of an event when the distance between any two points is greater than δ can be obtained in the form of the product $P_{[a,b]}(k,\delta)P_{[c,d]}(n-k,\delta)$. Therefore finally we can write:

$$P_{[a,b]\cup[c,d]}(n,\delta) = \sum_{k=0}^{n} P_{[a,b]}(k,\delta)P_{[c,d]}(n-k,\delta) \\ C_n^k \frac{(b-a)^k (d-c)^{n-k}}{\left[(d-c)+(b-a)\right]^n} \quad (15)$$

It should be noted, that $P_{[a,b]}(0,\delta)$ can be considered as equal to 1.

By summing all above-mentioned, from (11), (14) and (15), we obtain:

$$\int_{a}^{b} \left|[a,b]\setminus Uz(\delta)\right|^{n-1} P_{|[a,b]\setminus Uz(\delta)|}(n-1,\delta)dz = \\ = 2\int_{a}^{a+\delta} (b-z-\delta)^{n-1} P_{[z+\delta,b]}(n-1,\delta)dz + \\ + \int_{a+\delta}^{b-\delta} \sum_{k=0}^{n-1} C_{n-1}^k (z-\delta-a)^{n-1} (b-z-\delta)^{n-k-1} \times \\ \times P_{[a,z-\delta]}(k,\delta)P_{[z+\delta,b]}(n-k-1,\delta)dz \quad (16)$$

Thus, by considering (8) can be written the final recurrence relation:
when $b-a \le \delta$

$$P_{[a,b]}(0,\delta) = P_{[a,b]}(1,\delta) = 1, \\ P_{[a,b]}(n,\delta) = 0, \ \ n \ge 2 \quad (17)$$

when $\delta < b-a \le 2\delta$

$$P_{[a,b]}(0,\delta) = P_{[a,b]}(1,\delta) = 1,\ P_{[a,b]}(2,\delta) = \\ (b-a-\delta)^2 / (b-a)^2,\ P_{[a,b]}(n,\delta) = 0, \ \ n \ge 3 \quad (18)$$

when $b-a > 2\delta$

$$P_{[a,b]}(0,\delta) = P_{[a,b]}(1,\delta) = 1, \\ P_{[a,b]}(2,\delta) = (b-a-\delta)^2 \ / \ (b-a)^2, \\ P_{[a,b]}(n,\delta) = [1/(b-a)^n] \times \\ \times \{2\int_{\alpha}^{a-\delta} (b-z-\delta)^{n-1} P_{[z+\delta,b]}(n-1,\delta)dz + \\ + \int_{a+\delta}^{b-\delta} \sum_{k=0}^{n-1} C_{n-1}^k (z-\delta-a)^k (b-z-\delta)^{n-k-1} \times \\ \times P_{[a,z-\delta]}(k,\delta)P_{[z+\delta,b]}(n-k-1,\delta)dz\} \quad (19)$$

4. Results of the study. The set of equalities (17)-(19) allows us to carry out the procedure of consistent finding $P_{[0,1]}(n,\delta)$ for different n.

Software for the described method of finding probability of not merging biological objects depend-

ing on their number and system resolution was developed. That allowed make experimental research. The results of calculations are presented in fig. 1.

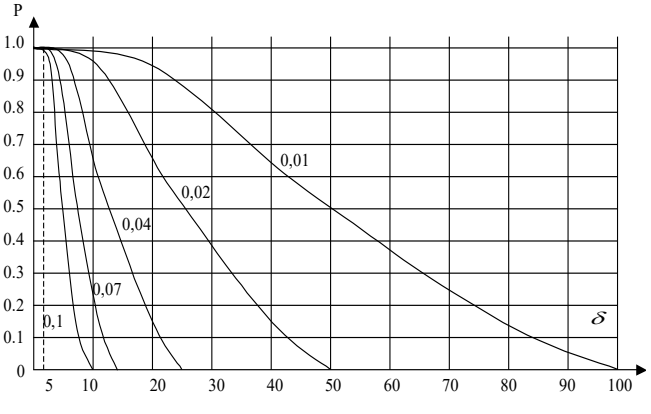

Fig.1. Plots of the probability $P_{[0,l]}(n,\delta)$ for different values δ from 0.01 to 0.1.

5. Conclusion.

The approach, proposed in the paper, allows determining the probabilities of the event, when two or more microbiological objects will register by the system as separated objects. The considered recurrence relations are obtained for different cases: when an interval in which objects hit does not exceed a value δ – the system resolution; when the size of an interval is between δ and 2δ; when the size of an interval is greater than the value 2δ. The obtained plots of the object recognition probability versus the system resolution values can be used to develop an optoelectronic systems or in the process of their application.

References

1. Dascanio J.J. Computer-Assisted Sperm Analysis / J.J. Dascanio. – Equine Reproductive Procedures, 2014. – 368 p.

2. Amann R.P. Computer-assisted sperm analysis (CASA): capabilities and potential developments / R. P. Amann, D. Waberski // Theriogenology. – 2014. – Т. 81. – № 1. – С. 5 - 17.

3. Pearl J. Probabilistic reasoning in intelligent systems: networks of plausible inference / J. Pearl. – Morgan Kaufmann, 2014. – 552 p.

4. Багиров В.А. Фертильность сперматозоидов и состояние хроматина: методы контроля (обзор) / В.А. Багиров, В.П. Кононов, Б.А. Иолчиев, П.М. Кленовицкий, Л.К. Эрнст // Сельскохозяйственная биология. – 2012. – №. 2. – С. 3 -13.

5. Шигимага В.А. Исследование проводимости клеток при изменении осмотической концентрации среды / В.А. Шигимага, Ю.Е. Мегель // Eastern-European Journal of Enterprise Technologies. – Kharkov. – 2011. – № 2/5(50). – С. 53 - 55.

6. Megel Y.E. Boosting membrane transmissibility with an impulse electric field / Yu.E. Megel, V.A. Shigimaga, S.N. Kovalenko, T.B. Belikova // Engineering of nature management. – 2016. – №. 2 (6). – С. 119 -124.

7. Megel Yu. Automation of measurement of objects geometrical parameters / Yu. Megel, I. Kalimanova, A. Rybalka, S. Kovalenko, S. Kovalenko // 27th International Scientific Symposium Metrology and Metrology Assurance, September 8 -12, 2017, Sozopol, Bulgaria, 2017. – P. 255 - 259.

8. Megel Y. E. Influence of Frequency Fluctuations on Accuracy of Measurement of Biomaterial Parameters in Nondestructive Testing Systems / Y. E. Megel, V. A. Muntyan // Telecommunications and Radio Engineering. – 2007. – Т. 66. – №. 6. – P. 571 - 575.

9. Megel Y. Operations Research, Calculus of Variations and Optimal Control, Part II / Y. Megel, A. Rudenko, S. Kovalenko, M. Tarulli, G. Venkov. – TU-Sofia, 2016. – 188 p.

10. Вентцель Е. С. Теория вероятностей. 10-е изд./ Е. С. Вентцель. – Москва: Высшая школа, 2006. – 575 с.

Анотація

Метод оцінки параметрів автоматизованої системи для аналізу рухливості мікробіологічних об'єктів

Ю. Є. Мегель, А. І. Рибалка, Г. П. Вєнков, С. В. Коваленко, І. В. Чалий

Велика кількість біологічних об'єктів характеризується рухливістю. Залежність напрямку руху окремих клітин або їх колонії при використанні зовнішнього світлового, електромагнітного опромінення або хімічного впливу на розчин може бути досліджена за допомогою оптико-електронної системи аналізу треків і підрахунку кількості об'єктів. У разі аналізу характеристик таких об'єктів їх число в полі зору системи є випадковою величиною. Аналогічна проблем, виникає в біотехнології відтворення та селекції сільськогосподарських тварин при створенні нових порід з запрограмованими властивостями і при застосуванні методу штучного запліднення.

У процесі створення таких систем виникає задача визначення ймовірності, з якою система може розрізняти два або більше об'єктів, що з'являються в її полі зору. У статті пропонується метод визначення ймовірності в залежності від роздільної здатності системи і кількості об'єктів, що знаходяться в її полі зору.

Так як число об'єктів в полі зору може варіюватися або задаватися шляхом зміни концентрації розчину, метою роботи є визначення числа об'єктів, при якому не відбувається злиття об'єктів, з ймовірністю, яка перевершує деякий наперед заданий рівень значущості, досить близький до одиниці.

Ключові слова: *мікробіологічний об'єкт, оптико-електрона система, роздільна здатність, поле зору, ймовірність*.

Аннотация

Метод оценки параметров автоматизированной системы для анализа подвижности микробиологических объектов

Ю. Е. Мегель, А. И. Рыбалка, Г. П. Венков, С. В. Коваленко, И. В. Чалый

Большое число биологических объектов характеризуется движением. Зависимость направления движения отдельных клеток или их колонии при использовании внешнего светового, электромагнитного облучения или химического воздействия на раствор может быть исследована с помощью оптико-электронной системы анализа треков и подсчета количества объектов. В случае анализа характеристик таких объектов их число в поле зрения системы является величиной случайной. Аналогичная проблем, возникает в биотехнологии воспроизводства и селекции сельскохозяйственных животных при создании новых пород с запрограммированными свойствами и применении метода искусственного осеменения.

В процессе создания таких систем возникает задача определения вероятности, с которой система может различать два или более появляющихся в ее поле зрения объектов. В статье предлагается метод определения вероятности в зависимости от разрешающей способности системы и количества объектов, находящихся в ее поле зрения.

Так как число объектов в поле зрения может варьироваться или задаваться за счет изменения концентрации раствора, целью работы является определение числа объектов, при котором не происходит слияния объектов, с вероятностью, превосходящей некоторый наперёд заданный уровень значимости, достаточно близкий к единице.

Ключевые слова*: микробиологический объект, оптико-электронная система, разрешающая способность, поле зрения, вероятность*.

Представлено від редакції: В.Ф Ужик / Presented on editorial: V.F Uzhik
Рецензент: О.Ю. Сергієнко / Reviewer: O.Yu. Sergiyenko
Подано до редакції / Received: 26.12.2017